International Tax Leadership Among Asymmetric Countries

Jean HINDRIKS
Yukihiro NISHIMURA
International Tax Leadership among Asymmetric Countries

Jean HINDRIKS\textsuperscript{1} and Yukihiro NISHIMURA\textsuperscript{2}

June 29, 2014

Abstract

Multinational companies can shift profit and income between branches in order to reduce the overall tax liabilities of the company. The result is a tax competition between countries. In this paper we consider the sequential choice of tax rates to illustrate the potential effects of tax leadership. We use a profit shifting model with multinational firms that operate in two countries, large and small. Governments compete by setting source-based corporate income taxes. We show that: (i) the sequential tax equilibria always Pareto dominate the simultaneous tax equilibrium. (ii) Each country prefers to follow than to lead the tax game. (iii) The tax leadership by the large country risk-dominates the tax leadership by the small country. Therefore our analysis provides a plausible explanation for the endogenous emergence of the tax leadership by the large countries. The results are contrasting with previous results in the literature.

\textbf{JEL Classification:} C72, F23, F68, H25, H87.

\textbf{Keywords:} Profit shifting, Tax competition, Endogenous timing, Second-mover advantage, Risk-dominance

1 Introduction

Globalization has prompted the emergence of multinational firms with divisions in different countries. It posed challenges for tax authorities across

\textsuperscript{*}This paper was presented at the Workshop in CORE (February 10, 2014) and University of Liège (February 14, 2014). We thank participants for helpful comments.

\textsuperscript{1}\textsuperscript{1}CORE, Université catholique de Louvain. jean.hindriks@uclouvain.be

\textsuperscript{2}\textsuperscript{2}Graduate School of Economics, Osaka University. ynishimu@econ.osaka-u.ac.jp
the world. Indeed, the widely used source-based taxation creates an incentive for multinational companies to shift profit and income between branches in order to reduce the overall tax liabilities of the company (Devereux and Griffith (2003) and Auerbach et al. (2010)). As a result countries that previously exploited their positions as leaders in setting tax rates by imposing high corporate tax rates increasingly found themselves competing with other jurisdictions to attract mobile investment and profits. In this context, small countries are generally thought to face the most elastic corporate tax bases, and therefore to have the strongest incentives to offer low corporate tax rates.\(^1\) Whether countries actually design their policies based on these assumed elasticities is another matter. Hines (2005) finds that in 1982, small countries had an average tax rate of 38.9 percent, while the average for large countries was 43.7 percent. In 1999, the average small country tax rate was 31.1 percent, and the average large country tax rate was 33.8 percent. Progressive reduction of the effects of country sizes on corporate tax rates is one of the implications of intensified international tax competition, since it is the ability to exploit market power that permits large countries to benefit from higher tax rates.

The presence of profit shifting is not so easy to detect, but it is getting some convincing empirical support. Mintz and Smart (2003) in a study of Canadian firms distinguishes between firms that operate in multiple jurisdictions which may engage in income shift, and firms that operate in a single jurisdiction. According to their preferred estimate, the elasticity of taxable income with respect to corporate income tax rates for “income shifting” firms is 4.9, compared with 2.3 for other, comparable firms.\(^2\)

Confronting the profit shifting, the arm’s length principle can sound attractive, but in reality, adjustments at most result in a poor approximation. Alternative solution to the transfer pricing regulation is to use the formula

\(^1\)Theoretically, there are possible mitigating factors such as strategic behavior and distortions induced by other policies. While there are few tests of the proposition that the supply of capital to small countries is more elastic than the supply of capital to large countries, they do not matter that much, because, in most models, it follows as an implication of the relatively small domestic corporate tax bases in small countries.

\(^2\)Bartelsmann and Beetsma (2001) find that more than 65 percent of the additional revenue resulting from a unilateral tax increase is lost because of income shifting even among countries. Clausing (2003) and Swenson (2001) find that taxation has a significant effect on intra-firm prices. IMF (2011) uses Bankscope data on banks to find that 1 percent-point higher tax rate reduces reported fiscal profits by between 6 and 8.5 percent.

\(^3\)Adjustments based on the arm’s length principle can lead to double taxation if the tax authorities in two jurisdictions do not agree on the price to be charged. Moreover, transactions within the group are specific with no equivalent market price, and the valuation of intangibles and royalty payments between related branches is a real challenge for
apportionment (FA) rule as in the USA and Canada for the allocation of
income across states. However, FA requires significant international coordi-
nation among the countries where a particular multinational company has
operations, since each country would like to pick the distribution rule that is
the most favorable to itself.\textsuperscript{4}

The main obstacle to international coordination is the absence of global
institutional structures. In a recent paper, Hindriks et al. (2013) propose
a solution based on a system of voluntary tax sharing among asymmetric
countries. They also contrast this voluntary tax sharing agreement with the
standard fiscal equalization scheme. In this paper, we suggest the possible
emergence of international tax leadership as a form of commitment on tax
setting. In our framework, the equilibrium outcome is the tax leadership by
large (core) countries in which smaller (periphery) countries would respond
in setting their own taxes. This can be viewed as a form of partial coordi-
nation where there is no need of international tax agency. This approach
is somewhat related to Cooper’s (1986) argument that in price competition
setting, firms may develop devices, such as the most-favored customer price
clause, to coordinate on a preferable price leadership outcome with symme-
try among firms. We seek to extend this argument to tax competition with
asymmetry between tax jurisdictions. A relevant device to coordinate on the
tax leadership outcome could then be some form of most-favored country tax
clause.

In most tax competition models, it is assumed that countries set taxes
simultaneously. But Schelling (1960) pointed out that the viability of the
simultaneous-move equilibrium becomes dubious once countries’ commit-
ment is considered. Some countries can rely on their relative size to take the
leadership in the tax competition game, with smaller size countries following.
There is a small literature in which tax leadership by the large countries is
assumed but not demonstrated.\textsuperscript{5} On the other hand, Hamilton and Slutsky
(1990) dealt with the issue of endogenous timing by using pre-play stage.
This approach has been mostly used in industrial organization, notably by
Amir and Stepanova (2006), who developed the endogenous timing game in
a Bertrand duopoly game with strategic complementarities in prices. They

\textsuperscript{4}The EU’s attempt to bring about an FA union is, at best, a half-measure because even
if small numbers of countries in which multinational companies are operating are left out
of this FA union, the problem of transfer pricing will persist.

\textsuperscript{5}Wang (1999) assumed that the larger country is the Stackelberg leader, and compared
the outcome with the Nash outcome. Baldwin and Krugman (2004) used the model with
economies of agglomeration to explain why tax rates remain higher in the core country
than in the periphery, by assuming that the core country moves first.
obtained the result that the strong (low-cost) firm is the leader if the cost asymmetry is high enough.\textsuperscript{6} More recently Kempf and Rota-Graziosi (2010, 2011) have adopted this approach in tax competition models, respectively in a capital mobility model, and in a cross-border shopping model. Their main finding is that under some restrictions on the countries asymmetry, the small country would lead in the equilibrium of endogenous timing game.

We will revisit Kempf and Rota-Graziosi’s (2010, 2011) results using instead a profit mobility model (instead of capital mobility model or cross border shopping model). We will use the model developed by Hindriks et al. (2013) with two countries of different market sizes and two multinational firms with branches in each country. We show that the Nash equilibrium in taxes is Pareto dominated by the Stackelberg equilibria in taxes. We also show that each country has a second-mover advantage no matter the difference in market sizes. Lastly, we show that the leadership by the large country is preferable (in the risk dominance sense) to the leadership by the small country, which is the opposite of the Kempf and Rota-Graziosi (2010,2011). We attribute the difference to their implicit assumption on the tax response elasticity.

2 The model

The model used follows closely Hindriks et al. (2013). There are two countries and two multinationals. The two countries are denoted by 1 and 2. A homogeneous good is produced. The market size in each country is characterized by the following linear (inverse) demands

\[ p_1(q_1) = \gamma_1 - \beta q_1 \quad \text{and} \quad p_2(q_2) = \gamma_2 - \beta q_2, \]

where different market size is represented by \( \gamma_1 > \gamma_2 \), so that country 1 is the large country with a higher demand for the good. This is equivalent to say that countries differ in population size (obviously alternative interpretation could be that countries differ in income per capita with the rich country 1 displaying higher willingness to pay for the good).

Two multinational firms, \( a \) and \( b \), have branches in each country and compete à la Cournot in each national market. The unit production cost is normalized to zero; so is the cost of shipping goods across countries. Multinational firms may, at some cost, shift profits between branches so as to minimize their total tax liability. Let \( \pi^j_i = p_i(q^a_i + q^b_i)q^j_i \) be the profit effectively generated by firm \( j = a, b \) in country \( i = 1, 2 \), associated with

\textsuperscript{6}On the other hand, in Cournot duopoly model, Amir and Grilo (1999) obtained the opposite timing with the weak (high-cost) firm as the leader.
production decisions \((q^a_i, q^b_i), i = 1, 2\). The firm must decide how much profit to report in country \(i\), \(\tilde{\pi}^j_i\), subject the constraint that total reported profit equals total realized profit \(\tilde{\pi}^j_i + \tilde{\pi}^j_2 = \pi^j_1 + \pi^j_2\). The cost of profit shifting is convex and non-fiscally deductible with for \(i = 1, 2\) and \(j = a, b\)

\[
C(\pi^j_1, \tilde{\pi}^j_i) = 2\delta (\pi^j_1 - \tilde{\pi}^j_i)^2,
\]

where \(\delta\) is a scaling cost parameter.\(^7\) This may reflect the cost of hiring accounting experts, the expected fine to be paid to the government, or the expected market sanction when caught cheating on tax liabilities. The parameter \(\delta\) captures the intensity of competition. The cost of profit shifting is also independent of the direction of shifting (outward and inward shifting are cost-equivalent).

Government \(i\) sets a source-based tax rate \(t_i\) on the profit reported within its jurisdiction by multinational firms. Tax revenue in country \(i\) is

\[
R_i = t_i (\tilde{\pi}^a_i + \tilde{\pi}^b_i) = t_i \tilde{\pi}_i.
\]

We assume that governments seek to maximize their fiscal revenue.\(^8\) Adding the consumer surplus in the governmental objective function will not affect the analysis, because in this profit shifting model consumer surplus is independent of the tax choices. We assume that \(t_i \leq 1\), for \(i = 1, 2\).

The sequence of events is as follows. First, both countries choose their tax rates so as to maximize their tax revenue. Second, given tax choices, multinational firms compete à la Cournot on each local market and choose a level of production in each country and the fraction of profit to be shifted to low-tax jurisdiction.

Proceeding backwards, we analyze firms’ decisions in each country, given the tax choices made at the previous stage. Firms determine the quantities to produce in each market and the amount of profit shifting. The problem of firm \(a\) is to maximize

\[
(1 - t_1)\tilde{\pi}^a_1 + (1 - t_2)\tilde{\pi}^a_2 - 2\delta(\pi^a_1 - \tilde{\pi}^a_i)^2,
\]

subject to \(\tilde{\pi}^a_1 + \tilde{\pi}^a_2 = p_1(q^a_1 + q^b_1)q^a_1 + p_2(q^a_2 + q^b_2)q^a_2\), with the inverse demand function given by (1). Hindriks et al. (2013) showed the following:

\[
q^a_i = q^b_i = \gamma_i/(3\beta), \quad p_i = \gamma_i/3, \quad \pi^a_i = \pi^b_i = \gamma_i^2/(9\beta), \quad i = 1, 2.
\]

\(^7\)See also Nielsen et al. (2005), Peralta et al. (2006), Amerighi and Peralta (2010), and Swenson (2001). See also Huizinga and Laeven (2008) for a slightly different specification.

\(^8\)See Hindriks et al. (2013) for a justification of this objective function.
Namely, productions and prices are not dependent on $t_i$’s.\(^9\) We can normalize production by assuming $\gamma_1 = \frac{3}{2} \sqrt{\beta (1 + \epsilon)}$, $\gamma_2 = \frac{3}{2} \sqrt{\beta (1 - \epsilon)}$ with $\epsilon \in (0, 1)$. Then minimizing each multinational firm’s total tax liabilities subject to the concealment cost gives the total profit reported in country $i$, $\tilde{\pi}_i = \tilde{\pi}_i^a + \tilde{\pi}_i^b$ in the equilibrium:

$$\tilde{\pi}_1 = \frac{1 + \epsilon}{2} - \frac{t_1 - t_2}{2\delta}, \quad \tilde{\pi}_2 = \frac{1 - \epsilon}{2} + \frac{t_1 - t_2}{2\delta}$$

(2)

The aggregate profit is therefore equal to 1 regardless of the tax choices, making the tax game a zero-sum game. Note that for identical taxes $t_1 = t_2$ the distribution of aggregate profits between two countries is solely determined by the market size parameter $\epsilon$, i.e., $\tilde{\pi}_1(t, t) = (1 + \epsilon)/2$ and $\tilde{\pi}_2(t, t) = (1 - \epsilon)/2$. Firms shift profits according to the tax difference $(t_1 - t_2)$ and the cost parameter $\delta$.

Notice that if countries cooperate, they would maximize the joint fiscal revenue

$$t_1 \left( \frac{1 + \epsilon}{2} - \frac{t_1 - t_2}{2\delta} \right) + t_2 \left( \frac{1 - \epsilon}{2} + \frac{t_1 - t_2}{2\delta} \right)$$

which leads to the cooperative optimal taxes $t_1^o = t_2^o = 1$, and maximal joint fiscal revenue equal to 1.

### 3 Tax competition

We now move to the first-stage game where both countries choose their tax rates. We will compare the simultaneous Nash equilibrium obtained in Hindriks et al. (2013) with the two possible Stackelberg equilibria, where either the large country leads, or the small country leads. We will then study the timing game in Section 4 by allowing the “pre-play” stage, in which the countries simultaneously and non-cooperatively decide whether to move “early” or “late”, leading to one of the three tax-games. This will determine which of these three tax games can emerge as a subgame perfect equilibrium of the extended “pre-play” timing game.

#### 3.1 Simultaneous game

This game is already analyzed in Hindriks et al. (2013), and so we will give a brief account. At the simultaneous-move game, each country noncooperatively chooses its own tax rate. Given $t_2$, the government of country 1 chooses

\(^9\)Given the non-deductibility of profit-shifting cost, the firms’ production decisions are independent of tax rates. Therefore adding consumer surplus in the objective function will not affect the governmental choices of taxes.
$t_1$ to maximize

$$R_1(t_1, t_2) = t_1 \left( \frac{1 + \epsilon}{2} - \frac{t_1 - t_2}{2\delta} \right),$$

which yields the tax reaction function of country 1

$$\hat{t}_1(t_2) = \delta \frac{1 + \epsilon}{2} + \frac{t_2}{2}. \quad (3)$$

Similarly, the tax reaction function of country 2 is

$$\hat{t}_2(t_1) = \delta \frac{1 - \epsilon}{2} + \frac{t_1}{2}. \quad (4)$$

It is interesting to note that the reaction functions differ only by the intercepts, with higher intercept for the large country 1. In our model, the large country has greater market power in the tax setting, which is related to the vertical intercept of the reaction function. This is similar to Amir and Stepanova (2006) in which the firms’ cost asymmetry implies different markups. On the other hand, in the capital mobility model of Kempf and Rota-Graziosi (2010), country’s asymmetry translates into different slopes of the reaction functions. In fact, both our model and Kempf and Rota-Graziosi (2010) yield higher tax base elasticities by small countries which was thought to be a key parameter for the analysis. However we will show shortly that what matters for the optimal leadership choice is the slope of the tax reaction functions which we call tax response elasticities.\textsuperscript{10} Notice also that taxes are strategic complements which is due to the positive tax externality. The Nash equilibrium taxes are,

$$t_1^N = \delta \frac{3 + \epsilon}{3} \quad \text{and} \quad t_2^N = \delta \frac{3 - \epsilon}{3}.$$ 

To insure interior solution, we assume throughout the rest of the paper that the profit shifting problem is binding and restrains tax choices.

$$\delta \leq \bar{\delta} = \frac{2}{3 + \epsilon} < 1.$$

The tax rate difference $t_1^N - t_2^N = 2\delta\epsilon/3$ implies that both firms shift profits from the large to the small country. However, this profit shifting is

\textsuperscript{10}In Kempf and Rota-Graziosi (2010, Section 4), the production functions are given by

$$F_i(K_i) = (a_i - b_iK_i)K_i \quad (i = 1, 2)$$

with $b_2 > b_1$, so firms in (small) country 2 are less productive than in (large) country 1, which brings different tax base elasticities (their “capital elasticities”). On the other hand, the tax reaction function is given by $\hat{t}_i(t_j) = b_j(t_j + 2Kb_j)/(b_i + 2b_j)$. The large country displays a greater tax response elasticity ($\partial \hat{t}_1/\partial t_2 > \partial \hat{t}_2/\partial t_1$).
not enough to offset the market size effect, and the large country ends up with a larger tax base in equilibrium,

\[ \tilde{\pi}_1^N = \frac{1}{2} + \frac{\epsilon}{6}, \quad \tilde{\pi}_2^N = \frac{1}{2} - \frac{\epsilon}{6}. \]

With a larger tax base and a higher tax rate, the large country also ends up with higher fiscal revenue in equilibrium,

\[ R_{1}^N = \frac{\delta}{18} (3 + \epsilon)^2, \quad R_{2}^N = \frac{\delta}{18} (3 - \epsilon)^2. \]

### 3.2 Stackelberg games

Two Stackelberg games can be considered, depending on which country takes the leadership.

Begin with the case in which (large) country 1 leads. Applying backward induction, (small) country 2, as the follower, chooses \( t_2 \), given \( t_1 \). Anticipating the best response of country 2 in (4), country 1, as the Stackelberg leader, maximizes \( R_1(t_1, \hat{t}_2(t_1)) \) with respect to \( t_1 \). The first-order condition is\(^{11}\)

\[ \frac{dR_1(t_1, \hat{t}_2(t_1))}{dt_1} = \frac{1 + \epsilon}{2} - \frac{t_1 - \hat{t}_2(t_1)}{2\delta} + t_1 \left( \frac{-1}{2\delta} \left( 1 - \frac{1}{2} \right) \right) = 0. \]

which gives the equilibrium tax rate of the Stackelberg leader, denoted by \( t_1^L \), as

\[ t_1^L = \frac{3 + \epsilon}{2}. \]

Substituting into (4), country 2’s best tax response, denoted by \( t_2^F \), is

\[ t_2^F = \frac{5 - \epsilon}{4}. \]

The tax differential is \( t_1^L - t_2^F = \delta(1 + 3\epsilon)/4 \), so that firms shift profits from the large leading country to the small following country. The equilibrium distribution of tax bases is,

\[ \pi_1^L = \frac{3}{8} + \frac{\epsilon}{8}, \quad \pi_2^F = \frac{5}{8} - \frac{\epsilon}{8}. \]

Hence, contrarily to the Nash equilibrium, in this Stackelberg equilibrium the profit-shifting effect more than offsets the market size effect, so that the

\(^{11}\)Notice that the revenue function is concave in the tax rate, \( d^2R_1(t_1, \hat{t}_2(t_1))/dt_1^2 = -1/(2\delta) < 0. \)
large leading country ends up with a smaller tax base. This is the result of tax leadership that enables the small country to steal more profit from the leading country by undercutting tax. The tax revenues are

\[ R^L_1 = \frac{\delta}{16}(3 + \epsilon)^2, \quad R^F_2 = \frac{\delta}{32}(5 - \epsilon)^2 \]

There are offsetting effects between the tax-rates differences \((t^L_1 > t^F_2)\) and the tax-base differences \((\tilde{\pi}^L_1 < \tilde{\pi}^F_2)\). As a result, the larger country only ends up better off if the size asymmetry is large enough.

\[ R^L_1 \geq R^F_2 \iff \epsilon \geq \epsilon^* \approx 0.314. \quad (5) \]

When \(\epsilon\) is large enough, the market-size effect dominates the cost of tax leadership. This is the opposite when \(\epsilon\) is small enough. In particular for arbitrary small asymmetry, the tax leader is necessarily worse off.

The countries’ difference is limited to the parameter \(\epsilon\), thus the second Stackelberg equilibrium in which (the small) country 2 leads can be easily obtained by analogy. The equilibrium taxes, given by \((t^F_1, t^L_2)\) are

\[ t^F_1 = \delta \frac{5 + \epsilon}{4}, \quad t^L_2 = \delta \frac{3 - \epsilon}{2}. \]

With size symmetry, the follower will necessarily tax less. With size asymmetry, the tax difference is \(t^F_1 - t^L_2 = \delta (3\epsilon - 1)/4\), so that firms shift profits from the large to the small country if and only if \(\epsilon > 1/3\). When \(\epsilon\) is large enough, the market-size effect dominates the follower incentive to undercut tax. This is the opposite when \(\epsilon\) is small enough. The equilibrium tax bases distribution is,

\[ \tilde{\pi}^F_1 = \frac{5}{8} + \frac{\epsilon}{8}, \quad \tilde{\pi}^L_2 = \frac{3}{8} - \frac{\epsilon}{8}. \]

When the small country leads, both the advantage to follow and the market-size effect reinforce each other to increase the large country’s tax bases. Therefore large country always ends up with more revenue in equilibrium.

\[ R^F_1 = \frac{\delta}{32}(5 + \epsilon)^2, \quad R^L_2 = \frac{\delta}{16}(3 - \epsilon)^2 \]

\(R^F_1 > R^L_2\) for all \(\epsilon \in (0, 1)\).

Naturally, the cost of shifting profit measured by \(\delta\), affects tax levels in equilibrium: lower \(\delta\) exacerbates the tax competition between countries and reduces the equilibrium taxes and revenue. It is straightforward to conclude that if \(\delta < \tilde{\delta}\) the joint tax revenue generated by the competitive outcome is smaller than under the cooperative outcome. This allows us to state our first proposition.
Proposition 3.1  In the Nash equilibrium as well as the two Stackelberg equilibria, there is under-taxation and joint tax revenue is sub-optimal. If profit shifting becomes easier (i.e., if $\delta$ decreases) Nash and Stackelberg equilibrium taxes decrease, and so does joint tax revenue.

3.3 Comparison of the tax games

Lemma 1 The equilibrium tax rates are ranked as follows: (i) For all $\epsilon \in (0,1)$ and $\delta < \bar{\delta}$, $t_N^1 < t_F^1 < t_L^1$ and $t_N^2 < t_F^2 < t_L^2$. (ii) For all $\epsilon \in (0,1)$ and $\delta < \tilde{\delta}$, $t_N^1 > t_N^2$ and $t_L^1 > t_F^2$. (iii) $t_F^1 < t_L^2$ if and only if $\epsilon < 1/3$.

Given the strategic complementarity (due to positive externalities in the tax game), in any Stackelberg equilibrium, tax rates are higher than the rates obtained at the Nash equilibrium. This is a general implication of supermodular games in which both reaction functions are increasing (Milgrom and Roberts, 1990). The leader can induce a favorable increase in the tax rate of its rival by increasing its own tax rate. As a result, the tax leader is induced to tax above the Nash level, and due to the strategic complementarity, the follower will also tax above the Nash level. On the other hand, the Stackelberg follower tends to undercut the tax rate in order to expand its own tax base.

Bucovetsky (1991) and Wilson (1991) pointed out in the capital mobility model that small countries tax less. The same results hold in the profit mobility model with the Nash equilibrium ($t_N^1 > t_F^1$) and the Stackelberg equilibrium where the large country leads ($t_L^1 > t_F^2$). However, when the small country leads, $t_L^1 > t_F^2$ if (and only if) $\epsilon > 1/3$ and so the small country may end up taxing more in equilibrium when leading the tax competition game. A similar result is obtained in Kempf and Rota-Graziosi (2010).

We define the concept of “second-mover advantage” as follows:

Definition 3.1 Country $i$ has a second-mover advantage if $R_{iF}^N > R_{iL}^N$.

Comparing the payoffs derived, we establish the following proposition:

Proposition 3.2 The equilibrium tax revenues are ranked as in Figure 1: (i) For all $\epsilon \in (0,1)$ and $\delta < \bar{\delta}$, both countries have a second-mover advantage in tax competition. (ii) For all $\epsilon \in (0,1)$ and $\delta < \tilde{\delta}$, $R_N^1 > R_N^2$ and $R_F^1 > R_F^2$. (iii) $R_L^1 < R_F^2$ if and only if $\epsilon < \epsilon^*$.  

Start with the case of symmetry ($\epsilon = 0$). From Lemma 1, both countries tax more in the Stackelberg equilibrium than in the Nash equilibrium. Due
Figure 1: Tax Revenues
to positive externalities, both $R_i^L$ and $R_i^F$ are higher than $R_i^N$ ($i = 1, 2$). When country 1 leads, the follower’s revenue gap ($R_2^L - R_1^L$) is first positive but decreasing in $\epsilon$ for $\epsilon \leq \epsilon^*$, and then the gap becomes negative when $\epsilon$ is high (i.e., $R_2^L - R_1^L < 0$ for $\epsilon > \epsilon^*$). This suggests that the relative size can to some extent offset the second-mover advantage. However, this does not mean that the large country would not prefer to be the follower. Alternatively when country 2 leads, the follower’s revenue gap ($R_1^F - R_2^L$) is always positive and increases in $\epsilon$ for all $\epsilon \in (0, 1)$.

We now contrast our second-mover advantage result with several previous papers. Firstly, using a capital mobility model, Kempf and Rota-Graziosi (2010) obtain a first-mover advantage by the small country for sufficient productivity differences. In their model the country asymmetry (marginal productivity of capital) translates into difference in the slopes of the reaction functions (see earlier footnote). Therefore, tax leadership by the small country can induce a greater favorable reaction in the tax rate of the following large country, than the other way round. On the other hand, in our model the asymmetry in market size translates into different intercepts but the same slopes of the reaction functions in (3) and (4), and the leadership of the large country emerges from the fact that the large country has a stronger incentive to tax than the small one. It is worth noting that the different tax base elasticities, that have been obtained in most models of asymmetric tax competition, including ours, are logically independent of the tax response elasticities implicitly assumed in Kempf and Rota-Graziosi (2010). Secondly, Kempf and Rota-Graziosi (2011), building upon Kanbur and Keen (1993) and Wang’s (1999) cross-border shopping model, obtain a first-mover advantage for the small country when asymmetry is sufficient. The difference with our model is that they use asymmetry in population densities whereas we use asymmetry in population sizes. Compared to France, Belgium is smaller in population size, but larger in population densities. This is relevant distinction but again, the first-mover advantage they obtain is partly driven by the induced difference in the tax response elasticity of each country.

\[ \frac{\partial \hat{\pi}_1}{\partial t_1} = \frac{t_1}{\delta(1 + \epsilon) - t_1 + t_2} \]

\[ \frac{\partial \hat{\pi}_2}{\partial t_2} = \frac{t_2}{\delta(1 - \epsilon) + t_1 - t_2} \]

In each country. As in most models, this elasticity, evaluated at equal tax rates, is higher for the small country.

\[ \frac{\partial \hat{\pi}_1}{\partial t_1} = \frac{t_1}{\delta(1 + \epsilon) - t_1 + t_2} \]

\[ \frac{\partial \hat{\pi}_2}{\partial t_2} = \frac{t_2}{\delta(1 - \epsilon) + t_1 - t_2} \]

12 Tax base elasticities in our profit mobility model are $-\frac{\partial \hat{\pi}_1}{\partial t_1} = \frac{t_1}{\delta(1 + \epsilon) - t_1 + t_2}$ and $-\frac{\partial \hat{\pi}_2}{\partial t_2} = \frac{t_2}{\delta(1 - \epsilon) + t_1 - t_2}$ in each country. As in most models, this elasticity, evaluated at equal tax rates, is higher for the small country.

13 Strictly speaking, another factor for the small country’s first-mover advantage in Kempf and Rota-Graziosi (2011) is the shifts in the intercepts of the reaction functions, depending on the location of the cross-shoppers (Figure 2 of Kanbur and Keen, 1993). They favor the welfare level of the small leadership equilibrium compared with that of the large leadership equilibrium.

Interestingly enough the structure of our game is closer to Hvidt and Nielsen’s (2001)
in Amir and Stepanova’s (2006) Bertrand duopoly game, the strong (low-cost) firm has a first-mover advantage when cost asymmetry is high. In our model it is impossible even though the second-mover advantage decreases as asymmetry increases.

According to the benefit of smallness in the capital-tax competition literature (Bucovetsky, 1991), the small country has a higher payoff than the large country in the tax competition equilibrium. In our setup of profit-shifting, the small country has a higher payoff (fiscal revenue) only when the large country leads and the asymmetry is small enough. Under the Nash equilibrium and the Stackelberg equilibrium where the small country leads, the large country has a higher fiscal revenue than the small one in the competitive equilibrium.\footnote{However, as we shall see in Proposition 5.2, the small country gains when moving from cooperation to competition.}

## 4 A timing game

A timing game, as suggested in Hamilton and Slutsky (1990), is defined as follows. At the “pre-play” stage, the countries simultaneously and non-cooperatively decide whether to move “early” or “late”. After the timing choice of each country is announced, the relevant tax competition game studied in the previous sections is played. Namely, if both players choose to move early (strategy \textit{Early}) or late (strategy \textit{Late}), the following tax competition is the simultaneous move of the decision of the tax rates, which end up with the tax rates \((t_1^N, t_2^N)\) and the tax revenues \((R_1^N, R_2^N)\). If one country \(i\) chooses \textit{Early} and the other country \(j\) chooses \textit{Late}, the tax competition is made by the sequential decisions and ends up with the Stackelberg-equilibrium tax rates \((t_i^L, t_j^F)\) and the tax revenues \((R_i^L, R_j^F)\). The relevant payoff structure of the timing game is reduced to the following normal form:

<table>
<thead>
<tr>
<th>country 1 (large)</th>
<th>country 2 (small)</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{Early}</td>
<td>(R_1^N, R_2^N)</td>
</tr>
<tr>
<td>\textit{Late}</td>
<td>(R_1^L, R_2^L)</td>
</tr>
</tbody>
</table>

In the Appendix we show the following:

\textbf{Proposition 4.1} For all \(\epsilon \in (0, 1)\) and \(\delta < \bar{\delta}\); (i) The Subgame Perfect Equilibria (SPEs) of the timing game are the two Stackelberg Equilibria.
(ii) Moving sequentially instead of simultaneously is Pareto-superior for both countries.

Therefore, a pre-play announcement is a commitment device that makes countries better-off compared to the conventional benchmark of simultaneous-move equilibrium. This is a partial form of coordination.

Due to the multiplicity of the equilibria of the timing game, a coordination issue appears regarding the equilibrium selection. We follow Harsanyi and Selten (1988) and adopt their risk-dominance criterion to our setting as follows:

**Definition 4.1** An equilibrium risk-dominates another equilibrium when the former is less risky than the latter, that is the risk-dominant equilibrium is the one for which the product of the deviation losses of reverting to the Nash equilibrium is the largest.

In our framework, equilibrium (Early, Late) risk-dominates equilibrium (Late, Early) if the former is associated with the larger product of deviation losses. Specifically:

$$\Gamma \equiv (R_1^L - R_1^N)(R_2^F - R_2^N) - (R_1^E - R_1^N)(R_2^L - R_2^N) > 0$$

The value of $R_1^L - R_1^N$ represents the threat of reverting to the Nash equilibrium when country 1 acts as a leader. The corresponding value for country 2 is $R_2^L - R_2^N$. The bigger the former is compared with the latter, the more country 1 has to lose in case of deviation from pre-play choice (Early, Late). The same explanation applies for $R_2^F - R_2^N$ and $R_1^F - R_1^N$. In short, this criterion aims for minimizing the risk of coordination failure. In the Appendix we show the following:

**Proposition 4.2** For all $\epsilon \in (0, 1)$ and $\delta < \bar{\delta}$, the SPE where the large country leads risk-dominates the other equilibrium.

The timing-game and risk dominance allow for a simple and natural explanation for the endogenous emergence of the large country’s leadership in the international tax competition. The point is similar to von Stackelberg’s (1934) critique of Cournot’s duopoly that the large firm is the most natural leader of the price competition game. Since the first-mover advantage is not observed in either country by Proposition 3.2(i), a possible Pareto-dominating relationship between SPE’s obtained in Amir and Stepanova (2006) and Kempf and Rota-Graziosi (2010, 2011) cannot arise in our model.
It is worth noting that our result of risk-dominance is opposite to Kempf and Rota-Graziosi (2011) where countries differ by population densities. It is also opposite to Kempf and Rota-Graziosi (2010) where countries differ in capital productivities. In both models, both countries are worse off when the large country leads, so that the risk-dominance criterion supports the SPE where the small country leads. We can attribute this peculiar result to their implicit assumption of different tax response elasticities discussed after Proposition 3.2.

5 Some policy issues

5.1 Total revenue and fiscal equalization

In terms of the total tax revenue the fiscal revenue gap among countries, we can show the following:

**Proposition 5.1** For all $\epsilon \in (0, 1)$ and $\delta < \bar{\delta}$, (i) $R^L_1 + R^E_2 > R^E_1 + R^L_2 > R^N_1 + R^N_2$, and (ii) $|R^L_1 - R^E_2| < R^E_1 - R^L_2$.

The risk-dominating SPE is better both in terms of the sum of the tax revenues and the revenue equalization. Thus tax leadership by the large country helps to increase total tax revenue and to reduce the fiscal revenue gap among countries. The reason of the fiscal equalization effect is that the large country leadership benefits directly to the small country that gets the second-mover advantage. The total revenue is maximized with tax leadership by the large country because this induces the country with the larger tax base to tax more.

5.2 Benefit from cooperation and minimal taxation

Another question is whether countries would benefit from cooperation, when they both set the cooperative tax rates $t^c_1 = t^c_2 = 1$ and get the respective tax revenue, $R^c_1 = (1+\epsilon)/2, R^c_2 = (1-\epsilon)/2$. Comparing this cooperative outcome with the risk-dominating SPE outcome, we obtain the following result.

**Proposition 5.2** $R^c_1 > R^L_1$ for all $\epsilon \in (0, 1)$ and $\delta < \bar{\delta}$. For all $\epsilon$, there exists $\delta(\epsilon) < \bar{\delta}$ such that $R^c_2 < R^E_2$ for all $\delta \in (\delta(\epsilon), \bar{\delta})$. $\delta(\epsilon)$ is decreasing in $\epsilon$ with $\delta(1) = 0$.

The potential advantage of the small country is its lower tax rate, which allows it to attract a fraction of the large country’s tax base. With tax harmonization, this is no longer possible. Thus, unless the fiscal competition
is too intense (low $\delta$) or the country sizes are too close (low $\epsilon$), the small country prefers the competition outcome to the tax harmonization outcome.

Another issue is the potential benefit of imposing a minimal tax rate. In the cross-border transaction model à la Kanbur and Keen (1993) with a revenue maximizing government, imposing a minimum commodity tax rate has the striking feature that both the small and the large countries could gain. The large country becomes less constrained by the threat of outward tax shopping and thus it chooses higher tax rate. The small country also benefits because of the induced increase in the tax rate in the large country thereby increasing the inward tax shopping. However, this does not extend to a Stackelberg setting in which the large country is assumed to be the leader (see Wang 1999 and Hvidt and Nielsen 2001). Without assuming leadership by the large country, we can prove similar result in our model with the endogenous tax leadership by the large country. The proof of the result is available upon request. We do not reproduce it here because it is very similar to Hvidt and Nielsen (2001).

6 Conclusion

"Multi-speed" Europe is used to describe the method of enhanced cooperation in Europe whereby common objectives are pursued by a group of Member States both able and willing to advance, it being implied that the others will follow later. The Member States concerned can thus move forward at different speeds with a sense of leadership. In this broad context of multi-speed cooperation in Europe, we evaluate the scope for international tax leadership to confront the increasing difficulty member states face in taxing global corporations. With globalization, capital is becoming increasingly mobile, but globalization also makes it easier for multinational companies to shift income and profit across branches to minimize their total tax liabilities.

Some authors have recently shown that tax leadership is not only natural but could even be beneficial (which refers to a first-mover advantage). Wang (1999) asserts that “it is natural and conceivable that, in a real-world situation of tax setting, the large region moves first” (p. 974). This conjecture was recently turned on its head by Kemp and Rota-Graziosi (2010) who find in a capital mobility model that leadership by the small region is, in fact, the most likely outcome. In this paper, we build upon a profit mobility model, to demonstrate that sequential move always dominates simultaneous tax game. However, we find that tax leadership by the large is preferable to tax leadership by the small in terms of risk dominance. This finding is opposite to the one of Kemp and Rota-Graziosi (2010). We attribute the difference to
the implicit assumption they make on the tax response elasticity. In their model, the tax rate of the large country is more sensitive to the tax rate of the other country. Therefore, relative to a Nash equilibrium, tax leadership by the small can induce a greater favorable reaction in the tax rate of the following large country, than the other way round. In our model, the tax responsiveness of each country is the same, and the leadership of the large country emerges as a pure size effect.

Appendix

Proof of Lemma 1: (i) \( t^F_1 - t^N_1 = \delta \left( \frac{5 + \epsilon}{12} - \frac{3 + \epsilon}{3} \right) = \frac{5}{12}(3 - \epsilon) > 0 \), and \( t^F_1 - t^N_1 = \delta \left( \frac{3 + \epsilon}{2} - \frac{5 + \epsilon}{4} \right) = \delta \frac{3}{4}(1 + \epsilon) > 0 \). (ii) \( t^N_1 - t^N_2 = \frac{2\delta}{3} \epsilon > 0 \), \( t^F_1 - t^F_2 = \frac{\delta}{4}(1 + 3\epsilon) > 0 \). \( t^F_1 - t^F_2 = \frac{\delta}{4}(3\epsilon - 1) < 0 \) if and only if \( \epsilon < 1/3 \). Q.E.D.

Proof of Proposition 3.2: (i) \( R^F_1 - R^L_1 = \delta \left( \frac{1}{32}(5 + \epsilon)^2 - \frac{1}{16}(3 + \epsilon)^2 \right) = \frac{\delta}{32}(-1 + \epsilon)^2 + 8 > 0 \) and \( R^F_2 - R^L_2 = \delta \left( \frac{1}{32}(5 - \epsilon)^2 - \frac{1}{16}(3 - \epsilon)^2 \right) = \frac{\delta}{32}(-1 - \epsilon)^2 + 8 > 0 \). (ii) \( R^N_1 - R^N_2 = \delta \left( \frac{1}{18}(3 + \epsilon)^2 - \frac{1}{18}(3 - \epsilon)^2 \right) = \frac{2\delta}{3} \epsilon > 0 \), \( R^F_1 - R^L_2 = \frac{\delta}{32}(-(11 - \epsilon)^2 + 128) > 0 \), and \( R^F_1 - R^L_2 = \frac{\delta}{32}((11 + \epsilon)^2 - 128) \). For \( \epsilon \in (0, 1) \), the last equation derives \( R^F_1 - R^F_2 \geq 0 \iff \epsilon \geq -11 + 8\sqrt{2} \), which is (5). Q.E.D.

Proof of Proposition 4.1: If country 1 chooses Early, since \( R^F_2 - R^N_2 = \frac{\delta}{288}(3 + \epsilon)(27 - 7\epsilon) > 0 \), country 2 chooses Late. If country 2 chooses Late, since \( R^L_1 - R^N_1 = \frac{\delta}{144}(3 + \epsilon)^2 > 0 \), country 1 chooses Early. Therefore, the pair of (Early, Late) constitutes a subgame perfect equilibrium. In the same way, the pair of (Late, Early) constitutes a subgame perfect equilibrium. Simultaneous move is never an equilibrium. This completes part (i) of the proposition. Part (ii) comes from Figure 1. Q.E.D.

Proof of Proposition 4.2: \( \Gamma = \frac{\delta}{144}(3 + \epsilon)^2 * \frac{\delta}{288}(3 + \epsilon)(27 - 7\epsilon) - \frac{\delta}{288}(3 -
\[(27+7\epsilon)\frac{\delta}{144}(3-\epsilon)^2 = \frac{1}{144 \times 288}\delta^2((3 + \epsilon)^3(27 - 7\epsilon) - (3 - \epsilon)^3(27 + 7\epsilon)) = \frac{1}{576}\delta^2(15 - \epsilon^2) > 0. \quad Q.E.D.
\]

Proof of Proposition 5.1: \(R_1^F + R_2^F = \delta \left(\frac{1}{16}(3+\epsilon)^2 + \frac{1}{32}(5-\epsilon)^2\right) = \frac{1}{32}\delta(43 + 2\epsilon + 3\epsilon^2), R_1^L + R_2^L = \delta \left(\frac{1}{18}(3+\epsilon)^2 + \frac{1}{18}(3-\epsilon)^2\right) = \delta \left(1 + \frac{1}{9}\epsilon^2\right). \quad \text{Clearly } R_1^L + R_2^L > R_1^F + R_2^F, \text{ and } (R_1^F + R_2^F) - (R_1^L + R_2^L) = \delta \left(-\frac{5\delta}{288}\epsilon^2 - \frac{\delta}{16}\epsilon + \frac{11}{32}\right) = \frac{1}{288}\delta(5\epsilon + 33)(3 - \epsilon) > 0. \quad \text{As to the differences in the tax revenue, for } \epsilon < \epsilon^*, \quad R_2^F > R_1^L \text{ and } (R_1^F - R_1^L) - (R_2^F - R_2^F) = \frac{11}{8}\epsilon > 0, \text{ and for } \epsilon \geq \epsilon^*, \quad (R_1^F - R_2^F) - (R_1^L - R_2^L) = \frac{\delta}{16}(7 - \epsilon^2) > 0. \quad Q.E.D.

Proof of Proposition 5.2: For country 1, \(R_1^L - \frac{1 + \epsilon}{2} = \frac{1}{16} \delta(3 + \epsilon)^2 - 8(1 + \epsilon) \leq -\frac{1}{8}(1 + 3\epsilon) < 0 \text{ for all } \delta \leq \delta.
\]

For country 2,

\(R_2^F - \frac{1 - \epsilon}{2} = \frac{1}{32} \delta(5 - \epsilon)^2 - 16(1 - \epsilon) \leq 0 \iff \delta \leq \frac{16(1 - \epsilon)}{(5 - \epsilon)^2} \equiv \delta(\epsilon).
\]

\(\delta(\epsilon) < \delta \text{ for all } \epsilon, \text{ with } \delta(1) = 0. \quad Q.E.D.

References


